Historical Data for Analytical Tooling - v3 - r2

Domain Portfolio: Observations | Domain: Historical | Usage Classification: **Limited Availability**

Geography: Global **Attribution Required:** No **Attribution Requirements:** N/A

Overview

The "Historical Data for Analytical Tooling" API outputs training data for use by Data Science teams to perform Initial Data Exploration and train Machine Learning models.

HTTP Headers and Data Lifetime - Caching and Expiration

For details on appropriate header values as well as caching and expiration definitions, please see **The Weather Company Data | API Common Usage Guide**.

Details

Once models are trained using data provided by this API, Predictive Analytics is often performed using the forecast counterpart of this API, Forecast Data for Analytical Tooling. The data provided by this API consists of the Average, Max and Min values, for each of 7 daily “periods” (expressed in local time relative to the postalKey or geolocation supplied in the call), for each of the specified weather aspects (“products”), for each day in the date range specified in the call. See Appendix A: Daily Aggregation Periods for a visualization of these aggregations and the daily periods over which they occur.

New data arrives hourly, and min/max/avg values for a given “daypart” (daily “period”) become available as soon as there are sufficient hours of data to compute them. The lag time of data availability is approximately 2 hours. The data used in this API is composed of 11 products from the gCOD (gridded Currents on Demand) dataset. Datasets, and even weather products within datasets, have varying start dates, as can be seen in the Data Dictionary for Product IDs section of this document. Given that the dayparts need certain hours of data to compute their min/max/avg values, and that a given queried point will imply an offset from Zulu time, it is important to note that queries near the arrival of a product in a dataset may have incomplete dayparts which would not be present in the response.

A paging mechanism is employed to allow clients to fetch years of historical data while keeping individual HTTP transactions within a time period that Weather’s authentication mechanisms will allow. On the first request the **next** query parameter is omitted or empty. If the request results in too much data to be satisfied in a single HTTP transaction a **X-Next-Token** header is returned in the HTTP response. This token may be echoed to subsequent calls to the same URL to fetch subsequent data. When no data remains to be fetched the **X-Next-Token** is not set.

This API only supplies a CSV response. The CSV response is formatted such that the first row of data defines the columns for all subsequent values. It is up to the client to determine how to handle entries for which there was no data (e.g. if not all weather products were selected). See Appendix B: CSV Sample and Appendix C: CSV Column Titles.

The products endpoint can be used to discover valid values for the productId parameter to the data endpoint. The (case-insensitive) word “all” can be used to include all available products in the response.

The characteristics of the underlying data can be found in a subsection of the HoD Utility API description.

For periods where there is no data, the data will be represented by an empty field. For averaged time periods, the average is derived from the hourly samples that are collected and does not include periods where no data is collected. For more, see: Quick Reference and Best Practices.

URL Construction

Atomic API URL Examples:
Data Request by Geocode: Required Parameters: geocode, productId, language, units, format, apiKey, startDate, endDate | Optional Parameter: next
https://api.weather.com/v3/wx/observations/historical/analytical/r2?**geocode=39.86,-104.67&productId=Temperature&startDate=20170101&endDate=20170630**&language=en-US&units=s&**format=csv&apiKey=yourApiKey**
Data Request by Postal Key: Required Parameters: postalKey, productId, language, units, format, apiKey ,startDate, endDate, | Optional Parameter: next
https://api.weather.com/v3/wx/observations/historical/analytical/r2?**postalKey=81657:US&productId=Temperature&startDate=20170101&endDate=20170630**&language=en-US&units=s&**format=csv&apiKey=yourApiKey**
Product Discovery: Required Parameters: apiKey
https://api.weather.com/v3/wx/observations/historical/analytical/r2/products?**apiKey=yourApiKey**

Valid Parameter Definitions

Parameter Name	Valid Parameter Value	Description
postalKey	81657:US	Colon-delimited Postal code and Country Code combination
productId	Dewpoint, FeelsLike, Gust, MSLP, PrecipAmount, RelativeHumidity, SnowAmount, Temperature, UVIndex, Visibility, WindSpeed	Weather aspect for which aggregate values are to be retrieved. Can be individual (single) product, comma delimited list of multiple products, or the (case-insensitive) word “all” to get them all. Derived from /products endpoint
startDate	20180101	ISO8601 formatted date
endDate	20180101	ISO8601 formatted date
next	20180101 or empty	Page token
language	en-US	Only en-US supported at this time
units	s (SI: Système international d'unités)	The unit system in which numerical values should be expressed
format	csv	Output format = csv is valid

Data Dictionary for Product IDs

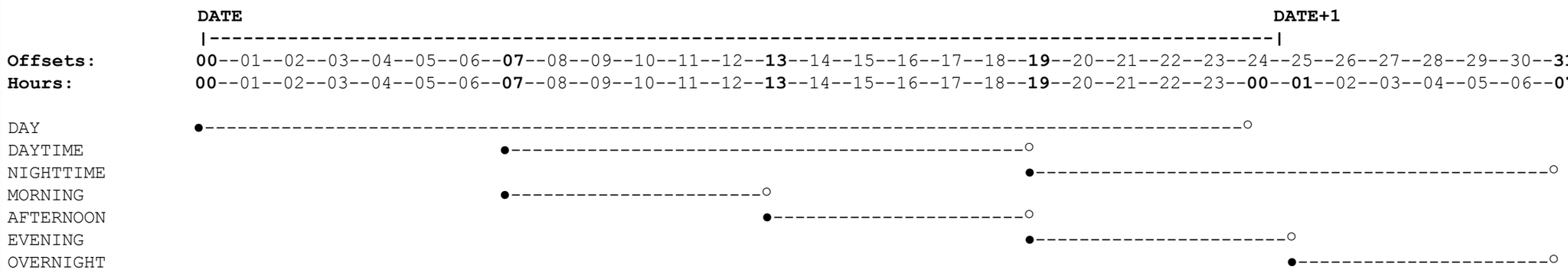
Weather Variable	Description	Unit
MSLP	Hourly mean sea level pressure. Archive begins from July 17, 2017	Pa
RelativeHumidity	The relative humidity of the air, which is defined as the ratio of the amount of water vapor in the air to the amount of vapor required to bring the air to saturation at a constant temperature. Relative humidity is always expressed as a percentage.	%
Temperature	The temperature of the air, measured by a thermometer 1.5 meters (4.5 feet) above the ground that is shaded from the other elements.	K
UVIndex	UV index 0 to 2 is Low, 3 to 5 is Moderate, 6 to 7 is High, 8 to 10 is Very High, 11 to 16 is Extreme	Enumerated Integer
Dewpoint	The temperature which air must be cooled at constant pressure to reach saturation. The Dew Point is also an indirect measure of the humidity of the air. The Dew Point will never exceed the Temperature. When the Dewpoint and Temperature are equal, clouds or fog will typically form. The closer the values of Temperature and Dew Point, the higher the relative humidity.	K
FeelsLike	feels like temperature. An apparent temperature. It represents what the air temperature "feels like" on exposed human skin due to the combined effect of the wind chill or heat index.	K
WindSpeed	The wind information reported in the hourly current conditions corresponds to a 10-minute average called the sustained wind speed. Sudden or brief variations in the wind speed are known as "wind gusts" and are reported in a separate data field (gust).	m/s
Visibility	The horizontal visibility at the observation point. Visibilities can be reported as fractional values particularly when visibility is less than 2 miles. Visibilities greater than 10 statute miles(16.1 kilometers) which are considered “unlimited” are reported as “999”; in your feed. You can also find visibility values that equal zero. This occurrence is not wrong. Dense fogs and heavy snows can produce values near zero. Fog, smoke, heavy rain and other weather phenomena can reduce visibility to near zero miles or kilometers.	m
Gust	The maximum expected wind gust speed.	m/s
PrecipAmount	Millimeters of precipitation occurring in the hour queried	mm
SnowAmount	Meters of snow occurring in the hour queried	m

HTTP Headers/Data Elements & Definitions (See below for full list of headers)

Header/Field Name	Description	Type	Range	Sample	Nulls Allowed
X-Next-Token	An opaque string that the caller of the HDAT api may include as the query parameter **next** to fetch data that could not be returned in an earlier data request. By “opaque” we mean that the string shall have no necessary meaning to the caller other than that it was returned by the HDAT service and may be echoed back to that service to continue data iteration. The empty string submitted as a request shall mean that the first page of data is desired. The empty string returned as a response shall mean that no further data is available.	[string]	Any valid HTTP header string.	20180101	Y

Appendix A: Daily Aggregation Periods

"DATE" = DATE the aggregations are referred to as being created for
Below is a depiction of the semantics that both Historical and Forecast Aggregations will use when considering the time ranges for that **"DATE"**



Appendix B: CSV Sample

All numerical data are decimal numbers.

date	postalKey	latitude	longitude	DewpointLocal DayAvg	DewpointLocal DayMax	DewpointLocal DayMin	DewpointLocal DaytimeAvg	DewpointLocal DaytimeMax	DewpointLocal DaytimeMin	DewpointLocal NighttimeAvg	DewpointLocal NighttimeMax	DewpointLocal NighttimeMin	DewpointLocal MorningAvg	DewpointLocal MorningMax	DewpointLocal MorningMin	DewpointLocal AfternoonAvg	DewpointLocal AfternoonMax	Dew fter
2018-09-19	53719:US	43.03095817	-89.51221458	63.875	66	62	63.75	65	62	65.58333333	66	65	62.83333333	63	62	64.66666667	65	
2018-09-20	53719:US	43.03095817	-89.51221458	68.125	71	65	69.41666667	71	66	66.91666667	69	62	68.83333333	70	66	70	71	
2018-09-21	53719:US	43.03095817	-89.51221458	54.95833333	68	45	52.33333333	60	46	44.41666667	47	42	56.16666667	60	53	48.5	52	
2018-09-22	53719:US	43.03095817	-89.51221458	45.83333333	49	42	46.33333333	48	43	48	49	47	45.16666667	47	43	47.5	48	
2018-09-23	53719:US	43.03095817	-89.51221458	52.08333333	56	47	53.16666667	56	47	54.16666667	56	52	51.16666667	54	47	55.16666667	56	
2018-09-24	53719:US	43.03095817	-89.51221458	56.33333333	60	52	56.75	59	52	60.08333333	61	59	55	58	52	58.5	59	
2018-09-25	53719:US	43.03095817	-89.51221458	59.08333333	61	56	58.75	60	56	58	59	56	59.66666667	60	59	57.83333333	60	
2018-09-26	53719:US	43.03095817	-89.51221458	54.29166667	59	49	53.83333333	59	49	48.75	51	46	57	59	55	50.66666667	53	
2018-09-27	53719:US	43.03095817	-89.51221458	46.25	49	43	45.66666667	48	43	45.33333333	46	44	47	48	45	44.33333333	46	
2018-09-28	53719:US	43.03095817	-89.51221458	44.79166667	46	43	44.75	46	43	43.16666667	45	41	45.5	46	43	44	45	
2018-09-29	53719:US	43.03095817	-89.51221458	42.20833333	44	40	42.41666667	44	40	41.16666667	42	40	43.33333333	44	41	41.5	43	
2018-09-30	53719:US	43.03095817	-89.51221458	41.70833333	43	40	41.83333333	43	40	42.66666667	43	42	42.16666667	43	40	41.5	42	
2018-10-01	53719:US	43.03095817	-89.51221458	43.25	45	42	43.58333333	45	42	43.08333333	44	42	44.16666667	45	42	43	44	
2018-10-02	53719:US	43.03095817	-89.51221458	44.29166667	46	42	45	46	43	43.83333333	45	42	45.16666667	46	43	44.83333333	46	

Appendix C: CSV Column Titles

The CSV columns are: date, postalKey, latitude, longitude, followed by the distribution of the requested weather aspects over 7 aggregation periods over 3 aggregation values (Avg, Max, Min). Note: All columns with the time period of 'LocalDay' record a 'day' from midnight to midnight. All other columns represent a 'day' as 7 am to 7 am. There will always be the 1st 5 columns of:

```
date, geoType, geoTypeId, latitude, longitude
```

followed by columns named by the "word product" of: `[Weather Variable] x "Local" x [Day Part Title] x ["Avg" | "Max" | "Min"]`

So, if you only specified productId=Dewpoint, the column titles of the resulting csv (following the 1st 5 above) would be:

```
DewpointLocalDayAvg,        DewpointLocalDayMax,        DewpointLocalDayMin,
DewpointLocalDaytimeAvg,    DewpointLocalDaytimeMax,    DewpointLocalDaytimeMin,
DewpointLocalNighttimeAvg,  DewpointLocalNighttimeMax,  DewpointLocalNighttimeMin,
DewpointLocalMorningAvg,    DewpointLocalMorningMax,    DewpointLocalMorningMin,
DewpointLocalAfternoonAvg,  DewpointLocalAfternoonMax,  DewpointLocalAfternoonMin,
DewpointLocalEveningAvg,    DewpointLocalEveningMax,    DewpointLocalEveningMin,
DewpointLocalOvernightAvg,  DewpointLocalOvernightMax,  DewpointLocalOvernightMin
```